DSSTox Log File:

EPA High Production Volume (HPV) Challenge Program (HPVCSI)

Structure-Index File

(last updated 18 February 2008)

Description: Information in this file documents creation, review, and update process for the DSSTox HPVCSI SDF file, and provides summary information on database content. The first section summarizes the process used for creating the initial DSSTox SDF files and the quality assurance checks and procedures employed. A table providing field and data counts offers summary overview of HPVCSI file content. The Log Table documents modifications and revisions to the database content or format in version updates. To obtain the most current version of this Log File and a record of any new modifications, or to report errors in this file, a user should consult the DSSTox HPVCSI SDF Download Page: http://www.epa.gov/ncct/dsstox/sdf_hpvcsi.html

QA and Development Notes for v1:
Data included in HPVCSI underwent a series of quality review checks prior to publication of initial launch version. An initial SDF for an earlier version of the HPV chemical lists was kindly provided by Jay Tunkel (Syracuse Research Corp) at the start of HPVCSI construction. This file was expanded to match the full chemical inventory content (CAS and chemical names) of the 3 published EPA HPV Challenge Chemical Lists (http://www.epa.gov/chemrtk/pubs/update/hpvchmlt.htm). The initial CAS and chemical names for the HPV Challenge Lists inventory were imported into Excel and cross-referenced to the existing DSSTox Master File chemical inventory, all issues pertaining to mismatch of CAS and chemical name were resolved, and DSSTox Standard Chemical Fields (including structures) were assigned when matches to existing DSSTox Master File inventory were found. For those chemical substances not in the existing DSSTox Master File inventory, we employed a number of commercial and public Internet resources for converting name to structure ((ACD/Name, version 9) or assigning chemical structures based on CAS, and populating DSSTox Standard Chemical Fields (for details of general QA review procedures, see http://www.epa.gov/ncct/dsstox/ChemicalInfQAProcedures.html). IUPAC systematic chemical names, **STRUCTURE_ChemicalName_IUPAC**, were computed using the ACD/Labs IUPAC Name-Generation software (ACD/Name, version 9); SMILES were generated with ACD/Labs ChemFolder (version 9). **InChI** codes were automatically generated from the final DSSTox SDF using the publicly available program, wINChI1.exe, downloadable from the NIST InChI website (http://www.iupac.org/inchi/).

Notes for v2a:
HPVCSI_v2a has no additional chemical substance records, but includes 5 new Source fields specifying chemical class, sponsorship and testing status of HPV Challenge Program substances obtained from downloadable tables on the HPV Challenge Program Source website (http://www.epa.gov/hpv/). In addition, URLs were obtained from EPA's Office of Pollution Prevention and Toxics (OPPT), with assistance from Matt Martin (EPA), pointing to 986 chemical-specific data pages on the new EPA HPV Information System (HPVIS) (http://www.epa.gov/hpvis/). HPVCSI_v2a additionally includes minor QA corrections, field entry revisions, field changes, etc. Changes to DSSTox Standard Chemical Fields include new ID fields: **DSSTox_RID**, **DSSTox_Generic_SID** and **DSSTox_FileID** (replacing **DSSTox_SID** and **DSSTox_ID_FileName** (see http://www.epa.gov/ncct/dsstox/MoreonStandardChemFields.html). Entries in the **TestSubstance_Description** field also have been simplified. This version includes URLs only for those chemical substances included on the original HPVCSI_v1a file, corresponding to the 3 published EPA HPV listings.

Notes for v2b:
Previously, in HPVCSI_v2a, URLs for chemical substance-specific toxicity test data pages in the EPA HPV Information System (HPV-IS) were included for approximately one third of the HPVCSI inventory. However, since the HPVCSI file is static and does not fully cover the content of the

HPV-IS, a new DSSTox file has been created to cover the HPV Information System Data (HPVISD) inventory more precisely, and to include the corresponding chemical data page URLs for the HPV-IS system. As a result, these URLs have been removed in this v2b update of HPVCSI.

In addition, with the availability of more detailed description of chemical mixture substance records in the published HPV-IS Robust Summaries, 14 new representative structures for mixtures have been added, with 2 previously classified mixtures converted to “single chemical compound” status, and based on more complete information, many more representative structures have been slightly modified in HPVCSI_v2b.

Notes for v2c: HPVCSI_v2c includes minor structure changes/modifications and the new **STRUCTURE_InChIKey** field (25 character abbreviated InChI for use in structure-indexing applications) added as a DSSTox Standard Chemical Field to all DSSTox files.

Source Website: EPA's High Production Volume (HPV) Challenge Program website - http://www.epa.gov/hpv/.

Log of SDF Modifications and Version/revision updates:

Date	DSSTox SDF File Name	Modifications from previous version	Additional Notes

10Apr2006	HPVCSI_v1a_3548_10Apr2006	Initial launch publication; no previous published versions.	HPVCSI is considered a "live" database meaning that further expansion of the database to include additional data is likely. Future updates also will correct reported errors provided by users or incorporate DSSTox format changes.
04Oct2006	HPVCSI_v1b_3548_04Oct2006	Revision includes updated URL links in the field **Website_URL** to the three published EPA HPV lists provided on the Source Website.	Numerous structure modifications and changes in stereochemical rendering throughout DSSTox data files following major quality review.
15Aug2007	HPVCSI_v2a_3548_15Aug2007	Revised Standard Fields: **DSSTox_SID** has been replaced by two new ID fields **DSSTox_RID** and **DSSTox_Generic_SID**. **DSSTox_ID_FileName** has been replaced by new ID field: **DSSTox_FileID**. Entries in **TestSubstance_Description** field have been simplified. Entries in **ChemicalNote** that pertained specifically to DBPCAN have been moved to Source-Specific field: **Note_DBPCAN** Previous contents of **Website_URL** field have been moved to field entitled: **HPVProgram_ChemicalList_URL** Chemical substance-specific URLs pointing to toxicity data in the HPV Information System have been added to modified **Website_URL** field for 986 records (20 URLs for substances not on one of the original 3 published HPV chemical lists were not included in this version).	Minor QA corrections throughout file. List 20 URLs and chemical substances in HPV-IS but not included in this file are provided in table below.
25Oct2007	HPVCSI_v2b_3548_25Oct2007	**Website_URL** field containing the chemical data pages in the EPA HPV Information System (HPV-IS) has been deleted and URLs moved to newly published HPVISD_v1a file, which focuses exclusively on the content of the HPV-IS. Based on additional information on mixture characteristics extracted from HPV-IS, 12 new representative structures have been added to HPVCSI_v2b.	
15February 2008	HPVCSI_v2c_3548_15Feb2008	56 structures were modified. New Standard Field added: **STRUCTURE_InChIKey**	All corrections or changes to structure information noted in **Note_HPVCSI** field, searchable by version (e.g., v2c).

HPVCSI SDF Content Summary:

HPVCSI SDF Content	Totals_v1	Totals_v2a	Totals_v2b	Totals_v2c
# Records	3548	3548	3548	3548
DSSTox Standard Chemical Fields	17	18	18	1819
HPVCSI Source Fields	1	6	6	6
Total # Fields	18	24	24	25
Chemical Content	**Counts_v1a**	**Counts_v2a**	**Counts_v2b**	**Counts_v2c**
STRUCTURE_ChemicalType:				
defined organic	2209	2209	2223	2226
inorganic	73	73	73	73
organometallic	181	181	181	181
no structure	1085	1085	1071	1068
STRUCTURE_TestedForm_ DefinedOrganic:				
parent	1911	1911	1921	1924
complex	61	64	63	63
salt	235	236	237	237
salt complex	2	2	2	2
TestSubstance_Description:				
single chemical compound	2008	1991	1991	1991
defined mixture or formulation	422	* (NA)	* (NA)	* (NA)
undefined mixture	1087	* (NA)	* (NA)	* (NA)
macromolecule	28	42	42	42
Unspecified or multiple forms	3	0	0	0
mixture or formulation	* (NA)	1515	1515	1518

** (NA) = field entry not applicable for DSSTox file version indicated*

Table of Excluded Chemical Substance URLs (v2a)
As of 15Aug2007, 20 chemical substances in the current HPV Information System (see table below) do not appear in the published EPA HPV Chemical Listings and, hence, are not included in HPVCSI_v2a. However, these new HPV chemicals will be incorporated into future version/revisions of the DSSTox HPVCSI file.

***** Update (25Oct2007) – v2b:*** All URLs have been removed from HPVCSI_v2b and these and the records in the table below have been added to the newly created DSSTox EPA HPV-Information System file (HPVISD).

HPVIS CASRN	HPVIS Chemical Name	HPVIS Website URL
84-75-3	1,2-Benzenedicarboxylic acid, dihexyl ester	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=100975
95-16-9	Benzothiazole	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101205
557-98-2	1-Propene, 2-chloro-	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=19288443
1918-00-9	Dicamba	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=156643
6891-44-7	Ethanaminium, N,N,N-trimethyl-2-[(2-methyl-1-oxo-2-propenyl)oxy]-, methyl sulfate	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101986
7173-51-5	1-Decanaminium, N-decyl-N,N-dimethyl-, chloride	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101225
13106-44-0	Ethanaminium, N,N,N-trimethyl-2-[(1-oxo-2-propenyl)oxy]-, methyl sulfate	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101981
13284-42-9	2 – Pentenenitrile	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101669
25103-52-0	Isooctanoic acid	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=100747
26760-64-5	Butene, 2-methyl-	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=100525
27070-58-2	Octadecene	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=100533
31565-23-8	di(tert-dodecyl) pentasulphide	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101988
50594-66-6	Benzoic acid, 5-[2-chloro-4-(trifluoromethyl)phenoxy]-2-nitro-	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=100940
61788-93-0	Amines, coco alkyldimethyl	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101259
61789-76-2	Amines, dicoco alkyl	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101256
61789-79-5	Amines, bis(hydrogenated tallow alkyl)	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101257
68425-15-0	Polysulfides, di-tert-dodecyl	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101990
68527-29-7	Tall oil, disproportionated, potassium salt	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101038
68583-56-2	tert-Dodecanethiol, sulfurized	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101989
68584-26-9	Benzenesulfonic acid, C10-16-alkyl derivs., magnesium salts	http://iaspub.epa.gov/oppthpv/quicksearch.display?pChem=101333